**SOUNDCLOUD DATA API – VISUALIZATION**

**Kyle Fyre, Taylor Lawrence, Varsha Teratipally**

**Abstract:**

Our project started out trying to tackle the problem of data on the platform Soundcloud. Currently, the information provided to content creators about their tracks and following is limited to essentially who played a track the most. With this project, our group wanted to explore the data provided by the Soundcloud API on a per user basis, to try to design new and creative visualizations to represent their following.

**Introduction:**

We thought it would be interesting to make a network visualization and see how many people either follow each other or are following the same people. Another visual we thought would be interesting, would be the number of times a song is remixed, so we would pick a set of songs and make a bubble graph with correlating to each different song, then the size of the bubble correlates to the amount of remixes that song has. Our team decided to focus on these attributes because there is so much data that flows into Soundcloud with how many profiles there are, we thought it would be interesting to see connections between profiles or artist. We decided to do the remix visual because Soundcloud is a very big place for users to make their own music or own remixes of songs, and we were interested in if we took a set of songs, how many remixes would there be of each.

**Related work:**

1. “Big Data meets research data from mobile diaries”

<https://ondeviceresearch.com/soundcloud-big-data>

Research goals:

Find the difference of existing users to mainstream audiences

Understand the wide market, attitudes, channels and competition

1. “Pandora Radio's Dominance Built On Big Data Edge”

<https://www.forbes.com/sites/amadoudiallo/2013/10/06/pandora-radios-dominance-built-on-big-data-edge/#60f2c6e5b595>

This article talks about how Pandora uses its data to track their listeners listening patterns to try and match each individual's music tastes.

1. “Spotify’s ‘music intelligence’ algorithm has identified some weird new music genres”

<https://www.dailydot.com/debug/spotify-weird-music-genres-echo-nest-algorithm/>

The article explains Spotify’s classification system, Echo Nest, that listens to music and tries to discover identifying elements and to become “trend aware”, then clusters songs together based on those trends to make different Radios, and playlists.

1. “Social Media Data Collection — Insights You Can Use”

<http://snaptrends.com/social-media-software/data-collection/>

This article highlights the importance of gathering social networking data. Soundcloud is not only a music app, but because users can follow, and communicate through the app, it stands as a type of social media. This article talks about why using social media data is a good thing, and how much insight it can give you.

1. “What Is the Facebook Algorithm?”

<https://www.bruceclay.com/blog/facebook-algorithm/>

This article highlights how facebook networks. It talks about how Facebook uses social interactions, engagements, and total number of followers, to create its networks.

1. “[ANALYSIS OF 30 YOUTUBE MUSIC VIDEO MASHUPS”](http://remixdata.net/2015/08/11/video-analysis-of-30-youtube-music-video-mashups/)

<http://remixdata.net/category/music-mashups/>

This is a study done that is very similar to our idea of looking at remixed songs and seeing how many there are based on a set of different songs. The research goal in this study was evaluate how remixed songs correlated to the originals and in what ways.

1. “The Rise Of Big Data And How Social Media Uses it”

<https://www.simplyzesty.com/blog/article/august-2012/the-rise-of-big-data-and-how-social-media-uses-it>

This article talks about how social media uses all its data, and the attributes they use to gather the data. For example, trends in likes, locations, friends, commenting patterns, usernames. This helps us in our study because it talks about the right things to use to gather data, when dealing with such a large dataset, like Soundcloud.

1. “AWS Case Study: SoundCloud”

<https://aws.amazon.com/solutions/case-studies/soundcloud/>

This study talks about how with the amounts of data flowing into Soundcloud every minute, how to deal with the large amounts of data and how to work with those large sets of data to get trends and such out of it, in order to continue Soundclouds networking aspect. This helps us because we are dealing with Soundcloud and having to deal with such large amounts of data, knowing how to work with the data can be really useful to us.

**Project Description:**

Networking visualization; shows different users and who they follow/ who follows them to try and see any correlations between followings.

**Features:**

1. The nodes are labelled with user account information.
2. Zooming can be done by moving the mouse from left to right or vice versa
3. A reset button which refreshes the dataset
4. When the user clicks on one node, the subgraph with its connected nodes will be displayed.
5. The size of the node is proportional to the number of users.

**Findings:**

Through the use of networks and graph theory, we were able to analyze the social structures of the sound cloud data api users. Density refers to the number of the followers where as the size of the node represents the popularity of the user.

**Technology:**  
1. HTML  
2. CSS  
3. JavaScript  
4. D3 library  
5. XAMPP for server  
  
**Steps To View The Visualization:**  
1. Clone the folder from the github repository into the local file system where the XAMPP server root files are located.  
2. Start the server using the XAMPP console.  
3. Open the web browser and browse the link "http://localhost/final-project-soundcloud-visualization-squad”.  
4. The visualizations will appear on the web page.